High Performance Algorithms and Software for Nonlinear Optimization, pp. 259-278
G. Di Pillo and A. Murli, Editors

Filter-type Algorithms for Solving Systems of Algebraic Equations and Inequalities

Roger Fletcher (fletcher@maths.dundee.ac.uk)

Sven Leyffer (sleyffer@maths.dundee.ac.uk)

Department of Mathematics
University of Dundee
Dundee DD1 4HN, Scotland, UK.

Abstract

The problem of solving a nonlinear system is transformed into a bi-objective nonlinear programming problem, which is then solved by a prototypical trust region filter SQP algorithm. The definition of the bi-objective problems is changed adaptively as the algorithm proceeds. The method permits the use of second order information and hence enables rapid local convergence to occur, which is particularly important for solving locally infeasible problems. A proof of global convergence is presented under mild assumptions.

Keywords: nonlinear systems, nonlinear programming, global convergence, filter, multiobjective optimization.

1 Introduction

The problem we consider is that of a general system of m nonlinear inequalities of the form

$$c_i(\mathbf{x}) \leq 0 \qquad i = 1, 2, \ldots, m, \qquad \mathbf{x} \in \mathbb{R}^n. \tag{1.1}$$

The formulation allows nonlinear equations to be included by expressing them as back-to-back inequalities, that is an equation $c(\mathbf{x}) = 0$ would be written as two separate inequalities $c(\mathbf{x}) \leq 0$ and $-c(\mathbf{x}) \leq 0$ in (1.1). We choose this form of notation because it allows us to consider relaxing one of these inequalities, but not the other.

We might first ask what we want from a method for solving (1.1). Of course we would like it to find a solution to given accuracy if one exists. Also we would like it to indicate when solutions do not exist. Then we would like the methods to be *reliable* (possibly by giving a global convergence proof), *efficient* (with an expectation of rapid local convergence), and *practical* (easy to implement, good numerical experience, solves large problems, etc.). In particular we observe that some methods for solving (1.1) can converge very slowly when solutions do not exist. We look for methods that aim to satisfy all the above criteria.

As we are solving nonlinear problems we can expect to use iterative methods, and we use $\mathbf{x}^{(k)}$, $k = 1, 2, \ldots$ to denote the successive iterates. The most obvious method to use is based on using successive linearization, in which a displacement $\mathbf{d}^{(k)}$ is calculated on iteration k that solves the system

$$c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} \leq 0 \qquad i = 1, 2, \ldots, m. \tag{1.2}$$

This calculation can be carried out using any method for Phase 1 of LP. Then $\mathbf{x}^{(k)}$ is updated to give $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{d}^{(k)}$. In the context of solving nonlinear equations, this is the well-known Newton's method. (We use the notation that $c_i^{(k)} = c_i(\mathbf{x}^{(k)})$, $\mathbf{a}_i = \boldsymbol{\nabla} c_i$, $\mathbf{a}_i^{(k)} = \mathbf{a}_i(\mathbf{x}^{(k)})$, and we let $A^{(k)}$ denote the matrix with columns $\mathbf{a}_i^{(k)}$ for $i = 1, 2, \ldots, m$). This method is known to exhibit local and second order convergence near a regular solution, but its global behaviour is unpredictable. It is also possible that the method can fail because the linearized system is inconsistent, even if (1.1) does have solutions.

An alternative approach is to pose (1.1) as a norm minimization problem

$$\underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad h(\mathbf{x}) \;:=\; \| \mathbf{c}^+(\mathbf{x}) \| \tag{1.3}$$

in some convenient norm $\| \cdot \|$, where $\mathbf{c}^+$ denotes the vector of infeasibilities $c_i^+ = \max(c_i, 0)$. This problem can also be solved by successive linearization (a so-called Gauss-Newton method), or else $h(\mathbf{x})$ might be used as a merit function in a Newton method with line search. The latter idea helps to improve the global properties of Newton's method, but there are still potential difficulties. Powell [5] gives an example in which convergence to a non-stationary point of $h(\mathbf{x})$ is observed, which is clearly unsatisfactory. Moreover, in the neighbourhood of a non-zero local minimum of $h(\mathbf{x})$, such as occurs when (1.1) is infeasible, very slow linear convergence can occur, due to the inadequacy of the linear model. We give an example of this below.

These deficiencies are addressed in the Sl_1QP trust region method in which the l_1 norm is used to define $h(\mathbf{x})$ and the l_1QP subproblem

$$\begin{aligned} \underset{\mathbf{d}\in\mathbf{R}^n}{\text{minimize}} \quad & \tfrac{1}{2}\mathbf{d}^T B^{(k)}\mathbf{d} + \sum_{i=1}^{m} (c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d})^+ \\ & \|\mathbf{d}\|_\infty \leq \rho, \end{aligned}$$

is solved on each iteration. The matrix $B^{(k)}$ approximates the Hessian of the Lagrangian and so enables the method to converge superlinearly when the problem is infeasible. Usual ideas for updating the trust region radius ρ can be employed. A disadvantage of the method is the relatively complicated form of the subproblem, which needs special purpose software for optimum efficiency. It is possible to convert the subproblem to a regular QP problem by adding extra variables, but this is less efficient and the resulting data structure is less convenient. Another disadvantage is that the method can sometimes be slow in following 'curved grooves' in $h(\mathbf{x})$. The method can be improved by allowing non-monotonic steps, for example by using second order correction (SOC) steps.

This paper has some ideas in common with the Sl_1QP method, but the main difference is to replace the norm minimization problem by a bi-objective optimization problem, so that the idea of a 'filter' can be used, which has proved very effective in solving NLP problems. Roughly speaking, the idea is to divide the constraints into two sets, indexed by J and $J^\perp$ respectively, where $J^\perp$ denotes the complement $\{1, 2, \ldots, m\} \backslash J$. One set ($J^\perp$) represents those constraints which are close to being satisfied, or for which the linearized constraint provides a good local model. Constraints in J are those that are proving difficult to satisfy, so that we are content to try to reduce a measure of their infeasibility. Thus we consider solving a so-called *nonlinear feasibility problem*

$$NFP(J) \left\{ \begin{array}{lll} \underset{\mathbf{x} \in \mathbf{R}^n}{\text{minimize}} & \displaystyle\sum_{i \in J} c_i(\mathbf{x})^+ & \\ \text{subject to} & c_i(\mathbf{x}) \leq 0 & i \in J^\perp \end{array} \right.$$

which is the minimization of the sum of constraint violations in J, subject to the system of inequalities given by constraints in $J^\perp$. As the constraints in J are infeasible, we require the inclusion

$$J \subseteq \mathcal{V}(\mathbf{x}) \tag{1.4}$$

to hold, where

$$\mathcal{V}(\mathbf{x}) = \{i \mid c_i(\mathbf{x}) > 0\} \qquad \text{and} \qquad \mathcal{A}(\mathbf{x}) = \{i \mid c_i(\mathbf{x}) = 0\}$$

denote respectively the sets of *violated constraints* and *active constraints* at $\mathbf{x}$. The sets J and $J^\perp$ are updated as the algorithm proceeds. The two objectives are to minimize the l_1 norm of constraint violations for (i) the J constraints, and (ii) the $J^\perp$ constraints. A similar idea is used by Fletcher and Leyffer [2] with good practical experience, but without a proof of global convergence. The main aim of this paper is to investigate to what extent recent developments in convergence theory for NLP filter methods can be applied in the context of solving (1.1). This leads us to propose a prototype filter method for solving (1.1) that is possibly more soundly based than that in [2].

We motivate our use of $NFP(J)$ with a few observations relating to the detection of infeasibility in (1.1). There are two situations which we might regard as giving some indication that (1.1) is infeasible, namely (1) a local minimizer of $h(\mathbf{x}) > 0$ is found, and (2) a point is located at which the linearized system (1.2) is infeasible, which we might describe as a point of local infeasibility. It is readily proved that (1) $\Rightarrow$ (2), so that finding a local minimizer of $h(\mathbf{x})$ (in any norm) provides a method for finding a point of local infeasibility. Of course, making a global statement about the infeasibility of (1.1) is impractical for nonlinear systems of any size, as it is equivalent to the global minimization of $h(\mathbf{x})$.

Solving $NFP(J)$ provides another way of finding a point of local infeasibility. (It is, in fact, equivalent to finding the minimizer of a scaled l_1 norm of $\mathbf{c}^+(\mathbf{x})$, showing that there is a close relationship between the solution of $NFP(J)$ and the use of the l_1 norm.) To see this, consider the HIMMELBD test problem in the CUTE test set. This has just two nonlinear equations

$$\begin{aligned} c_1(\mathbf{x}) &= x_1^2 + 12x_2 - 1 = 0 \\ c_2(\mathbf{x}) &= 49x_1^2 + 49x_2^2 + 84x_1 + 2324x_2 - 681 = 0 \end{aligned}$$

representing a parabola and a circle, respectively. The problem has two solutions in the vicinity of $(\pm 20, -35)$. Moreover, the two curves almost intersect at a point close to the origin, with the result that there are nonzero local minima of $h(\mathbf{x})$ in this vicinity. A local minimizer in the l_1 norm is at $(0.286, 0.279)$ and this is a solution of $NFP(J)$ for $J = \{1\}$ and $J^\perp = \{2\}$. This point is on the line $7x_2 + 124 = 36/x_1$ which contains all points of local infeasibility. Local minimizers of $h(\mathbf{x})$ in other common norms also lie on this line, close to the l_1 solution. However, different from all these is the solution of $NFP(J)$ for $J = \{2\}$ and $J^\perp = \{1\}$ at $(0.289, 0.076)$, which also provides a point of local infeasibility near the origin. This illustrates that, as regards finding a point of local infeasibility, solving $NFP(J)$ is equally effective as minimizing an overall norm function, and is convenient in that it allows us to formulate (1.1) as a bi-objective optimization problem. A plot of the parabola and circle near the origin shows two curves that are close to being parallel straight lines, illustrating the difficulty of finding a point of local infeasibility using only linear information. Posing the problem as $NFP(J)$, in either of the above ways, enables the feasibility problem to be treated as an NLP problem and facilitates the inclusion of second order information to give rapid convergence.

In Section 2 we describe a basic algorithm format involving a QP subproblem, and in Section 3 we prove global convergence to a point that satisfies Kuhn-Tucker (KT) conditions for $NFP(J)$, subject to an MFCQ constraint qualification. In Section 4 we discuss various ways in which a practical code may be developed, based on the prototype algorithm.

The constraints of (1.1) may include some linear constraints, including simple bounds on the variables, and the methods we consider are such as to maintain feasibility with respect to the linear constraints. Thus we assume that the linear con-

straints have been checked for consistency using any method for Phase 1 of LP. If the linear constraints are consistent, then this procedure yields an initial point that is feasible for the linear constraints, and if not the system can be rejected without further calculation.

2 An SQP Filter-type Algorithm

In this section we describe a prototype SQP filter trust region algorithm for which we can prove global convergence. This algorithm is flexible as regards practical implementation, and various decisions are open to choice. The algorithm is based on the SQP method applied to $NFP(J_k)$, where the set J_k is chosen adaptively as the algorithm proceeds. Thus we attempt to find a vector $\mathbf{d}$ which solves a QP subproblem

$$QP(\mathbf{x},\rho,J)\begin{cases} \underset{\mathbf{d}\in\mathbb{R}^n}{\text{minimize}} & \frac{1}{2}\mathbf{d}^T B\mathbf{d} + \sum_{i\in J}(c_i + \mathbf{a}_i^T\mathbf{d}) & \\ \text{subject to} & c_i + \mathbf{a}_i^T\mathbf{d} \le 0 & i \in J^\perp \\ & \|\mathbf{d}\|_\infty \le \rho. & \end{cases}$$

We require that J is such that $c_i + \mathbf{a}_i^T\mathbf{d} > 0$ for all $i \in J$, and also that $J \subseteq \mathcal{V}(\mathbf{x})$. Because our algorithm respects linear constraints, a consequence is that J is composed only of nonlinear constraints. Exactly how the set J is chosen is described below, although it is readily possible to achieve these requirements. We allow the possibility that J is empty, in which case, if the trust region constraint is inactive, the step $\mathbf{d}$ may be regarded as a Newton-step in a method for solving (1.1). The symmetric matrix B may be regarded as an approximation to the Hessian of a Lagrangian function, and is important in practice as it enables rapid convergence to be obtained when second order effects are significant. However, in our global convergence theory, B has only a minor role, so we do not reflect the fact that the $QP(\mathbf{x},\rho,J)$ depends on B in the notation.

We now turn to the definition of an NLP filter as introduced in [2]. In an NLP context, there are two conflicting aims, namely to minimize some objective function $f(\mathbf{x})$, and to satisfy the constraints, which we can regard as the minimization of some measure of infeasibility $h(\mathbf{c}(\mathbf{x}))$. In a filter pairs of values (h, f) are considered, obtained by evaluating $h(\mathbf{c}(\mathbf{x}))$ and $f(\mathbf{x})$ for various values of $\mathbf{x}$. A pair $(h^{(i)}, f^{(i)})$ obtained on iteration i is said to *dominate* another pair $(h^{(j)}, f^{(j)})$ if and only if both $h^{(i)} \le h^{(j)}$ and $f^{(i)} \le f^{(j)}$, indicating that the point $\mathbf{x}^{(i)}$ is at least as good as $\mathbf{x}^{(j)}$ in respect of both measures. The NLP filter is defined to be a list of pairs $(h^{(i)}, f^{(i)})$ such that no pair dominates any other. The notation $\mathcal{F}^{(k)}$ is used to denote the set of iteration indices j such that $(h^{(j)}, f^{(j)})$ is an entry in the current filter. (In practice it is not necessary to store the index set $\mathcal{F}^{(k)}$, the notation is just for theoretical convenience.) A point $\mathbf{x}$ is said to be "acceptable for inclusion in the filter" if its (h, f) pair is not dominated by any entry in the filter. This is the condition that

$$\text{either} \qquad h < h^{(j)} \qquad \text{or} \qquad f < f^{(j)} \tag{2.1}$$

for all $j \in \mathcal{F}^{(k)}$. We may also wish to "include a point $\mathbf{x}$ in the filter", by which we mean that its (h, f) pair is added to the list of pairs in the filter, and any pairs in the filter that are dominated by the new pair are removed. The filter is used as an alternative to a penalty function as a means of deciding whether or not to accept a new point in an NLP algorithm.

As in previous work it is necessary to slightly strengthen the inequalities (2.1) in order to force sufficient improvement in at least one of the measures of infeasibility. This modification enables a convergence proof to be made, but has negligible effect on practical performance. Here we use the type of test suggested originally by Chin and Fletcher [1] and used by Fletcher, Leyffer and Toint [3]. Thus the condition for a point being acceptable to the filter is that its (h, f) pair satisfies

$$\text{either} \qquad h \leq \beta h^{(j)} \qquad \text{or} \qquad f + \gamma h \leq f^{(j)} \tag{2.2}$$

for all $j \in \mathcal{F}^{(k)}$. Here β and γ are preset parameters such that $1 > \beta > \gamma > 0$, with β close to 1 and γ close to zero. This filter test has an important inclusion property that if a new point is added to the filter, the set of unacceptable points for the new filter always includes the set of unacceptable points for the old filter.

In the context of solving a feasibility problem, we shall use a similar notation to specify the two objective functions to be used in the filter algorithm, that is

$$f_J(\mathbf{c}) = \sum_{i \in J} c_i^+, \qquad h_J(\mathbf{c}) = \sum_{i \in J^\perp} c_i^+. \tag{2.3}$$

To simplify the notation, we denote $f_{J_k}^{(k)}$ by $f^{(k)}$, and $h_{J_k}^{(k)}$ by $h^{(k)}$. We also denote $\boldsymbol{\nabla} f_J(\mathbf{c}(\mathbf{x}))$ by $\mathbf{g}_J(\mathbf{x})$, or for example by $\mathbf{g}_J^\circ$ when $\mathbf{x} = \mathbf{x}^\circ$. At first sight one might think that a different filter would be needed for each set J that is generated by the algorithm. In fact little is lost by having a single filter in which pairs based on different sets J are entered. It is readily shown that if the same vector $\mathbf{c}$ is measured by $f_J(\mathbf{c})$ and $h_J(\mathbf{c})$ for various different sets J then none of the resulting pairs will dominate any other (excluding ties), and all can coexist in the filter.

The algorithm that we suggest is a trust region algorithm, and there are two important conditions that need to be satisfied if a global convergence proof is to be obtained. One is that the trial step $\mathbf{d}$ should be the optimal solution (or nearly so) of some model subproblem, and another is that if the nonlinearities are negligible then sufficient agreement between the actual and predicted reductions in the objective function should be obtained. How to obtain these conditions when the set J changes is a novel feature of the algorithm that we propose. The difficulty lies in the fact that changes to J change the objective and constraints in the underlying problem $NFP(J)$. In regard to $QP(\mathbf{x}^{(k)}, \rho, J)$, we introduce the terminology that the solution $\mathbf{d}$ and the set J *conform* if and only if $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} > 0$ for all $i \in J$. We shall see that if we choose J so that $\mathbf{d}$ and J conform then we are able to satisfy the above conditions.

We state our algorithm by means of the flow diagram of Figure 1. Associated with each point $\mathbf{x}^{(k)}$ at the start of an iteration is a set $J_k \subseteq \mathcal{V}_k$ that contains no linear constraint indices. There is also a current filter $\mathcal{F}^{(k)}$, and the current pair $(h^{(k)}, f^{(k)})$ is acceptable to $\mathcal{F}^{(k)}$ but is not included in it. The algorithm contains an inner loop in which the trust region radius ρ is successively reduced until a suitable new point is obtained. The inner loop is initialized with any value of $\rho \geq \rho_{\min}$, where $\rho_{\min} > 0$ is a preset parameter. For each value of ρ in the inner loop we determine a set $J_+ \subseteq J_k$ such that, when $QP(\mathbf{x}^{(k)}, \rho, J_+)$ is solved, then the solution $\mathbf{d}$ and the set J_+ conform. One way to do this is by means of the following loop.

1. Let $J = J_k$ and assume that a solution $\mathbf{d}$ to $QP(\mathbf{x}^{(k)}, \rho, J)$ exists.
2. If $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} > 0$ for all $i \in J$, then set $J_+ = J$ and exit.
3. Remove any indices i for which $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} \leq 0$ from J.
4. Find a solution $\mathbf{d}$ to $QP(\mathbf{x}^{(k)}, \rho, J)$ and goto step 2.

Note in steps 3 and 4 that the old $\mathbf{d}$ is feasible in the new QP subproblem, which must therefore be compatible. Only a finite number of repeats are required since J_k is a finite set and each repeat removes one or more indices from J. We now let $\mathbf{d}$ denote the solution of $QP(\mathbf{x}^{(k)}, \rho, J_+)$.

Next we calculate the vector $\mathbf{c}(\mathbf{x}^{(k)} + \mathbf{d})$ and determine another set $J_\oplus = J_k \cap \mathcal{V}(\mathbf{x}^{(k)} + \mathbf{d})$, which is the prospective value of J_{k+1} for the next iteration. The set $J_\oplus$ is obtained by deleting from J_k any indices i for which $c_i(\mathbf{x}^{(k)} + \mathbf{d}) \leq 0$. We note that the inclusions

$$J_+ \subseteq J_k \subseteq \mathcal{V}_k \qquad \text{and} \qquad J_\oplus \subseteq J_k \subseteq \mathcal{V}_k \tag{2.4}$$

both hold, but the sets J_+ and $J_\oplus$ may not always be the same. We use the set $J_\oplus$ to determine a candidate pair $(h_{J_\oplus}, f_{J_\oplus})$. This is tested for acceptability to the filter and to $(h^{(k)}, f^{(k)})$, and also possibly for sufficient reduction, as described below. If any of these tests fail then ρ is halved and the inner iteration is repeated. Otherwise the inner iteration terminates and the current values of ρ and $\mathbf{d}$ are designated respectively to be $\rho^{(k)}$, and $\mathbf{d}^{(k)}$. We then update $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{d}^{(k)}$ and assign $J_{k+1} = J_\oplus$, and proceed to the next iteration.

There is also the possibility to be considered that the constraints of $QP(\mathbf{x}^{(k)}, \rho, J_k)$ are inconsistent. In this case, any point $\mathbf{x}^{(k+1)} \in X$ and set $J_{k+1} \subseteq \mathcal{V}_{k+1}$ for which $(h^{(k+1)}, f^{(k+1)})$ is acceptable to the filter may be chosen. Such values may be obtained in many ways, for example by taking $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$ and choosing a subset of $\mathcal{V}_k$ for J_{k+1}. It is clear that acceptability to the filter can always be obtained by taking $J_{k+1} = \mathcal{V}_k$, since this provides $h^{(k+1)} = 0$. Note that $\mathcal{V}_k \backslash J_k$ cannot be empty as the constraints in the subproblem would all be linear and hence consistent. The choice of which subset to take is arbitrary, although in practice it seems better to take as small a subset as possible. This may be implemented by just relaxing a few constraints from $\mathcal{V}_k \backslash J_k$, possibly determined by using Phase 1 of an LP solver.

The tests in our filter algorithm are similar to those in [3], but with additional features to allow for the fact that J_k may change from iteration to iteration. If $q_J^{(k)}(\mathbf{d})$ refers to the objective function of $QP(\mathbf{x}^{(k)}, \rho, J)$, and $\mathbf{d}$ is the displacement obtained by solving $QP(\mathbf{x}^{(k)}, \rho, J_+)$, then we denote

$$\Delta q = q_{J_k}^{(k)}(\mathbf{0}) - q_{J_+}^{(k)}(\mathbf{d}) = f^{(k)} - q_{J_+}^{(k)}(\mathbf{d}) \tag{2.5}$$

as the *predicted reduction* in f_J, and

$$\Delta f = f^{(k)} - f_{J_\oplus}(\mathbf{c}(\mathbf{x}^{(k)} + \mathbf{d})) \tag{2.6}$$

as the *actual reduction* in f_J, where J_+ and $J_\oplus$ are defined as above. When the inner iteration terminates we designate the resulting values of (2.5) and (2.6) to be $\Delta q^{(k)}$ and $\Delta f^{(k)}$ respectively, and we have

$$\Delta f^{(k)} = f^{(k)} - f^{(k+1)} \tag{2.7}$$

by virtue of the method of choosing J_{k+1}.

We shall continue to use the terminology of [3] in which an *f-type step* is one in which an improvement in $f(\mathbf{x})$ is predicted, as defined by $\Delta q > 0$. In this case, we require a sufficient improvement condition

$$\Delta f \geq \sigma \Delta q \tag{2.8}$$

to hold, where $\sigma \in [0,\, 1]$ is a preset parameter. If the inner iteration terminates with an f-type step, we refer to the iteration as an *f-type iteration*. We note that $J_{k+1} \subseteq J_k$ for an f-type iteration. If $\Delta q \leq 0$, or if $QP(\mathbf{x}^{(k)}, \rho, J_k)$ is incompatible, then we refer to the step as an *h-type step*, and we expect an improvement in h to occur. We follow the ideas of [3] and only include the pair $(h^{(k)}, f^{(k)})$ in the filter if iteration k is an h-type iteration. We also use $\tau^{(k)}$ to denote the least value of $h^{(j)}$ for all $j \in \mathcal{F}^{(k)}$.

As in [2] and [3] we may also impose an upper bound u on $h^{(k)}$ to ensure that the linearizations of the active constraints are reasonably representative. This can be implemented by initializing the filter with an entry $(u/\beta, 0)$. We initialize $\mathbf{x}^{(1)}$ with any point in X, and choose J_1 such that $h^{(1)} \leq u$, if necessary by choosing $J_1 = \mathcal{V}_1$.

3 A Global Convergence Proof

In this section we present a proof of global convergence of the SQP-filter algorithm of Figure 1 when applied to (1.1). We make the following assumptions.

Standard Assumptions

1. All points $\mathbf{x}$ that are sampled by the algorithm lie in a non-empty closed and bounded set X.

2. The problem functions $c_i(\mathbf{x})$, $i = 1, 2, \ldots, m$ are twice continuously differentiable on an open set containing X.

3. There exists an $M > 0$ such that the Hessian matrices $B^{(k)}$ satisfy $\|B^{(k)}\|_2 \leq 2M$ for all k.

An additional assumption is also made later in the paper. It is a consequence of the standard assumptions that the Hessian matrices of the c_i are bounded on X and without loss of generality we may assume that they also satisfy bounds $\|\nabla^2 c_i(\mathbf{x})\|_2 \leq 2M$, $i = 1, 2, \ldots, m$. The assumptions on X are readily achieved by including simple upper and lower bounds in the constraint set. A consequence is that the sequence of iterates has at least one accumulation point, which we shall denote by $\mathbf{x}^\infty$. Quantities derived from $\mathbf{x}^\infty$ will be referred to, for example, by $\mathcal{V}_\infty$ or $\mathbf{c}^\infty$.

We would like our theory to make a meaningful statement when the constraint set contains some equations, which are represented as back-to-back inequalities as described in Section 1. We therefore define the set $\mathcal{E} \subseteq \{1, 2, \ldots, m\}$ as the index set that contains the inequalities arising from equations. Of course $\mathcal{E}$ has even cardinality. At a point $\mathbf{x}^\circ$ we shall denote $\mathcal{E}_\circ = \mathcal{E} \cap \mathcal{A}_\circ$ and we observe that the maximum rank of the set of vectors $\mathbf{a}_i^\circ$, $i \in \mathcal{E}_\circ$ is equal to $\frac{1}{2}|\mathcal{E}_\circ|$, on account of the duplication due to the presence of back-to-back inequalities. We then say that the Mangasarian-Fromowitz constraint qualification (MFCQ) holds at $\mathbf{x}^\circ$ for the problem $NFP(\mathcal{V}_\circ)$, if and only if both (i) the set of vectors $\mathbf{a}_i^\circ$, $i \in \mathcal{E}_\circ$ has rank $\frac{1}{2}|\mathcal{E}_\circ|$, and (ii) there exists a vector $\mathbf{s}$ that satisfies $\mathbf{s}^T\mathbf{a}_i^\circ = 0$, $i \in \mathcal{E}_\circ$ and $\mathbf{s}^T\mathbf{a}_i^\circ < 0$, $i \in \mathcal{A}_\circ \backslash \mathcal{E}_\circ$.

It is possible for the algorithm to terminate finitely, either if a feasible point of (1.1) is found, or if a KT point of a feasibility problem $NFP(\mathcal{V}_k)$ is found ($\mathbf{d} = \mathbf{0}$ solves $QP(\mathbf{x}^{(k)}, \rho, J_k)$ for some k (note that $J_k = \mathcal{V}_k$ is inferred)). Otherwise it follows that the sequence of iterates is infinite, and we prove in Theorem 1 below that there exists a limit point $\mathbf{x}^\infty$ that is either a feasible point of (1.1), or satisfies necessary conditions for $NFP(\mathcal{V}_\infty)$, if MFCQ holds at $\mathbf{x}^\infty$. In the context of solving $NFP(\mathcal{V}_\circ)$, we note that $\mathbf{x}^\circ$ is necessarily a feasible point of $NFP(\mathcal{V}_\circ)$, by definition of $\mathcal{V}(\mathbf{x})$. Thus, if MFCQ holds, KT necessary conditions for $\mathbf{x}^\circ$ to solve $NFP(\mathcal{V}_\circ)$ are equivalent to the statement that the set of directions

$$\begin{aligned} \{\mathbf{s} \quad | \quad & \mathbf{s}^T\mathbf{g}_{\mathcal{V}_\circ}^\circ < 0 && (3.1)\\ & \mathbf{s}^T\mathbf{a}_i^\circ = 0 \quad i \in \mathcal{E}_\circ && (3.2)\\ & \mathbf{s}^T\mathbf{a}_i^\circ < 0 \quad i \in \mathcal{A}_\circ \backslash \mathcal{E}_\circ \ \} && (3.3) \end{aligned}$$

is empty. This condition is therefore equivalent to the existence of KT multipliers (although we do not use this result in our proofs) and it has been shown (Gauvin [4]) that the multiplier set is bounded.

We first collect a number of useful results, analogous to those in [3]. First we state a result that is a consequence of the filter acceptance test.

Lemma 1 *Consider sequences* $\{h^{(k)}\}$ *and* $\{f^{(k)}\}$ *such that* $h^{(k)} \geq 0$ *and* $f^{(k)}$ *is monotonically decreasing and bounded below. Let constants* β *and* γ *satisfy* $0 < \gamma < \beta < 1$. *If for all* k

$$\text{either} \qquad h^{(k+1)} \leq \beta h^{(k)} \qquad \text{or} \qquad f^{(k)} - f^{(k+1)} \geq \gamma h^{(k+1)},$$

then $h^{(k)} \to 0$.

Corollary *Consider an infinite sequence of iterations on which* $(h^{(k)}, f^{(k)})$ *is entered into the filter, where* $h^{(k)} > 0$ *and* $\{f^{(k)}\}$ *is bounded below. It follows that* $h^{(k)} \to 0$.

Proof Both results are proved in [3].

Next we prove two simple lemmas that enable us to handle the second order terms in the analysis.

Lemma 2 *Consider minimizing a quadratic function* $\phi(\alpha)$ $(\mathbb{R} \to \mathbb{R})$ *on the interval* $\alpha \in [0, 1]$, *when* $\phi'(0) < 0$. *A necessary and sufficient condition for the minimizer to be at* $\alpha = 1$ *is* $\phi'' + \phi'(0) \leq 0$. *In this case it follows that* $\phi(0) - \phi(1) \geq -\frac{1}{2}\phi'(0)$.
Proof Trivial, but see [3].

Lemma 3 *Let the standard assumptions hold and let* $\mathbf{d}$, J_+ *and* $J_\oplus$ *be determined as described in Section 2. It then follows that*

$$c_i(\mathbf{x}^{(k)} + \mathbf{d}) \leq \rho^2 nM \qquad i \in J_\oplus^\perp \tag{3.4}$$

and

$$\Delta f \geq \Delta q - \rho^2 (m+1) nM. \tag{3.5}$$

Proof By Taylor's theorem, we may express

$$c_i(\mathbf{x}^{(k)} + \mathbf{d}) = c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} + \tfrac{1}{2}\mathbf{d}^T \nabla^2 c_i(\mathbf{y}_i)\mathbf{d} \tag{3.6}$$

where $\mathbf{y}_i$ denotes some point on the line segment from $\mathbf{x}^{(k)}$ to $\mathbf{x}^{(k)} + \mathbf{d}$. The set $J_\oplus^\perp$ can be divided into two sets, $J_k^\perp$ and $J_k \backslash J_\oplus$. For $i \in J_k^\perp$, we have by virtue of $J_k^\perp \subseteq J_+^\perp$ that $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} \leq 0$. Then (3.4) follows from (3.6), using the inequality $\|\mathbf{d}\|_2^2 \leq n\|\mathbf{d}\|_\infty^2$ and the bounds on $\mathbf{d}$ and $\nabla^2 c_i$. For $i \in J_k \backslash J_\oplus$, we have $c_i(\mathbf{x}^{(k)} + \mathbf{d}) \leq 0$ and (3.4) follows trivially.

It follows from (2.5), (2.6) and the properties of J_+ and $J_\oplus$ that

$$\Delta f - \Delta q = \tfrac{1}{2}\mathbf{d}^T B^{(k)}\mathbf{d} + \sum_{i \in J_+} (c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d}) - \sum_{i \in J_\oplus} c_i(\mathbf{x}^{(k)} + \mathbf{d}).$$

The quadratic term is bounded below by $-\rho^2 nM$ by virtue of the inequality $\|\mathbf{d}\|_2^2 \leq n\|\mathbf{d}\|_\infty^2$ and the bounds on $\mathbf{d}$ and $B^{(k)}$. For $i \in J_+ \cap J_\oplus$, the terms under the summation

are together equal to $-\frac{1}{2}\mathbf{d}^T\nabla^2 c_i(\mathbf{y}_i)\mathbf{d}$, by virtue of (3.6), and hence bounded below by $-\rho^2 nM$. For $i \in J_+\backslash(J_+ \cap J_\oplus)$, it follows by conformity that $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} > 0$. Finally, for $i \in J_\oplus\backslash(J_+ \cap J_\oplus)$ we have $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} \leq 0$ and hence, as in the first part of the lemma,

$$c_i(\mathbf{x}^{(k)} + \mathbf{d}) \leq \tfrac{1}{2}\mathbf{d}^T\nabla^2 c_i(\mathbf{y}_i)\mathbf{d} \leq \rho^2 nM.$$

Putting all these cases together yields (3.5). *q.e.d.*

The next lemma provides a condition on ρ which ensures that the QP step is acceptable to the filter entry for which $h^{(j)} = \tau^{(k)}$, and hence to all entries in $\mathcal{F}^{(k)}$.

Lemma 4 *Let the standard assumptions hold and let* $\mathbf{d}$, J_+ *and* $J_\oplus$ *be determined as described in Section 2. It follows that* $h_{J_\oplus}(\mathbf{c}(\mathbf{x}^{(k)} + \mathbf{d}))$ *satisfies the test* $h_{J_\oplus}(\mathbf{c}(\mathbf{x}^{(k)} + \mathbf{d})) \leq \beta\tau^{(k)}$ *if* $\rho^2 \leq \beta\tau^{(k)}/(mnM)$.

Proof It follows from (2.3) and (3.4) that $h_{J_\oplus}(\mathbf{c}(\mathbf{x}^{(k)} + \mathbf{d})) \leq \rho^2 mnM$, and the result follows if $\rho^2 \leq \beta\tau^{(k)}/(mnM)$. *q.e.d.*

Next we derive conditions on ρ under which an f-type step is acceptable in a neighbourhood of a point at which MFCQ holds that is not a local solution of $NFP(\mathcal{V}_\circ)$.

Lemma 5 *Let the standard assumptions hold and let* $\mathbf{x}^\circ \in X$ *be any point at which* $\mathcal{V}_\circ$ *is not empty and MFCQ holds, that is not a KT point of* $NFP(\mathcal{V}_\circ)$. *For any* $\mathbf{x} \in X$, *any* $J \subseteq \mathcal{V}(\mathbf{x})$, *and any* B *such that* $\|B\|_2 \leq 2M$, *let* $\mathbf{d}$, J_+ *and* $J_\oplus$ *be determined as described in Section 2. Then there exists a neighbourhood* $\mathcal{N}^\circ$ *of* $\mathbf{x}^\circ$ *and positive constants* μ *and* κ *such that if* $\mathbf{x} \in \mathcal{N}^\circ \cap X$, *if* J *is such that* $\mathcal{V}_\circ \subseteq J \subseteq \mathcal{V}_\circ \cup \mathcal{A}_\circ$, *and if* ρ *satisfies*

$$\mu h_{\mathcal{V}_\circ}(\mathbf{c}) \leq \rho \leq \kappa, \tag{3.7}$$

then it follows that $QP(\mathbf{x}, \rho, J)$ *is compatible and* $QP(\mathbf{x}, \rho, J_+)$ *has a feasible solution* $\mathbf{d}$ *at which the predicted reduction satisfies*

$$\Delta q \geq \tfrac{1}{3}\rho\varepsilon, \tag{3.8}$$

the sufficient reduction condition (2.8) holds, and

$$\Delta f \geq \gamma h_{J_\oplus}(\mathbf{c}(\mathbf{x} + \mathbf{d})), \tag{3.9}$$

which ensures acceptability to $(h^{(k)}, f^{(k)})$ *when* $\mathbf{x} = \mathbf{x}^{(k)}$ *and* $J = J_k$.

Proof Because $\mathbf{x}^\circ$ is not a KT point of $NFP(\mathcal{V}_\circ)$, and MFCQ holds, it follows that the vectors $\mathbf{a}_i^\circ$, $i \in \mathcal{E}_\circ$ have rank $\frac{1}{2}|\mathcal{E}_\circ|$ and there exists a vector $\mathbf{s}^\circ$ for which $\|\mathbf{s}^\circ\|_2 = 1$ that satisfies (3.1), (3.2) and (3.3). We note that these conditions imply that $\frac{1}{2}|\mathcal{E}_\circ| < n$. We let $A_\mathcal{E}$ denote a matrix with $\frac{1}{2}|\mathcal{E}_\circ|$ linearly independent columns chosen from the vectors $\mathbf{a}_i^\circ$, $i \in \mathcal{E}_\circ$, and we let $\mathbf{c}_\mathcal{E}$ denote the partition of $\mathbf{c}$ whose gradients are the columns of $A_\mathcal{E}$. By linear independence and continuity there exists

a neighbourhood of $\mathbf{x}^\circ$ in which $A_{\mathcal{E}}^+$ is bounded, where A^+ denotes $(A^TA)^{-1}A^T$. If $\mathcal{E}_\circ$ is not empty, we denote $\mathbf{p} = -A_{\mathcal{E}}^{+T}\mathbf{c}_{\mathcal{E}}$, which is the closest point to $\mathbf{d} = \mathbf{0}$ in the manifold of linearized active equality constraints, and let $p = \|\mathbf{p}\|_2$. Also we denote $\mathbf{s} = (I - A_{\mathcal{E}}A_{\mathcal{E}}^+)\mathbf{s}^\circ / \|(I - A_{\mathcal{E}}A_{\mathcal{E}}^+)\mathbf{s}^\circ\|_2$, which is the closest unit vector to $\mathbf{s}^\circ$ in the null space of $A_{\mathcal{E}}^T$. If $\mathcal{E}_\circ$ is empty, we set $\mathbf{p} = \mathbf{0}$, $p = 0$ and $\mathbf{s} = \mathbf{s}^\circ$. It follows from (3.1) and (3.3) by continuity that there exists a (smaller) neighbourhood $\mathcal{N}^\circ$ and a constant $\varepsilon > 0$ such that

$$\mathbf{s}^T\mathbf{g}_{\mathcal{V}_\circ} \leq -\varepsilon \qquad \text{and} \qquad \mathbf{s}^T\mathbf{a}_i \leq -\varepsilon, \quad i \in \mathcal{A}_\circ \backslash \mathcal{E}_\circ \tag{3.10}$$

when $\mathbf{g}_{\mathcal{V}_\circ}$, $\mathbf{a}_i$ and $\mathbf{s}$ are evaluated for any $\mathbf{x} \in \mathcal{N}^\circ$. By definition of $\mathbf{p}$, it follows that $p = O(h_{\mathcal{V}_\circ}(\mathbf{c}))$ so we can choose the constant μ in (3.7) sufficiently large so that $\rho > p$ for all $\mathbf{x} \in \mathcal{N}^\circ$.

First we derive some results about the QP subproblems that arise. From the trust region constraint $\|\mathbf{d}\| \leq \rho$ we can deduce the following. For inactive constraints $i \in (\mathcal{V}_\circ \cup \mathcal{A}_\circ)^\perp$ and $\mathbf{x} \in \mathcal{N}^\circ \cap X$, there exist positive constants $\bar{c}$ and $\bar{a}$, independent of $\mathbf{x}$, such that

$$c_i \leq -\bar{c} \qquad \text{and} \qquad \mathbf{a}_i^T\mathbf{s} \leq \bar{a},$$

where $c_i = c_i(\mathbf{x})$, etc., for all vectors $\mathbf{s}$ such that $\|\mathbf{s}\|_\infty \leq 1$. It follows that

$$c_i + \mathbf{a}_i^T\mathbf{d} \leq -\bar{c} + \rho\bar{a} \leq 0 \qquad i \in (\mathcal{V}_\circ \cup \mathcal{A}_\circ)^\perp. \tag{3.11}$$

if $\rho \leq \bar{c}/\bar{a}$. Thus inactive constraints at $\mathbf{x}^\circ$ remain inactive in the QP subproblem if κ in (3.7) satisfies $\kappa \leq \bar{c}/\bar{a}$. In a similar way, we can show that constraints $i \in \mathcal{V}_\circ$ remain infeasible if ρ is sufficiently small. Thus there exist positive constants $\hat{c}$ and $\hat{a}$, independent of $\mathbf{x}$, such that

$$c_i > \hat{c} \qquad \text{and} \qquad \mathbf{a}_i^T\mathbf{d} > -\rho\hat{a}.$$

Hence, if $\rho \leq \kappa \leq \hat{c}/\hat{a}$, it follows that $c_i + \mathbf{a}_i^T\mathbf{d} > 0$. Hence for any set J_+ that is determined, we deduce that $i \in J_+$, and hence $\mathcal{V}_\circ \subseteq J_+$. Moreover, it follows from the Taylor series (3.6) that

$$c_i(\mathbf{x} + \mathbf{d}) > \hat{c} - \rho\hat{a} - \rho^2 nM.$$

Thus, if $\rho \leq \kappa \leq \hat{c}/(\hat{a} + \hat{c}nM/\hat{a})$, it follows that $c_i(\mathbf{x} + \mathbf{d}) > 0$. Therefore, for any set $J_\oplus$ that is determined, we deduce that $i \in J_\oplus$, and hence $\mathcal{V}_\circ \subseteq J_\oplus$.

We now establish feasibility of the subproblem $QP(\mathbf{x}, \rho, J)$ for any $\mathbf{x} \in \mathcal{N}^\circ \cap X$ and any J for which $\mathcal{V}_\circ \subseteq J \subseteq \mathcal{V}_\circ \cup \mathcal{A}_\circ$. We consider the line segment defined by

$$\mathbf{d}^\alpha = \mathbf{p} + \alpha(\rho - p)\mathbf{s}, \quad \alpha \in [0, 1], \tag{3.12}$$

for a fixed value of $\rho > p$. Because the vectors $\mathbf{p}$ and $\mathbf{s}$ are orthogonal, it follows that

$$\|\mathbf{d}^\alpha\|_2 = \sqrt{p^2 + \alpha^2(\rho - p)^2} \leq \sqrt{\rho^2 - 2\rho p + 2p^2} \leq \rho$$

since $\rho > p$. Consequently $\|\mathbf{d}^\alpha\|_\infty \leq \rho$, and hence $\mathbf{d}^\alpha$ satisfies the trust region constraint for all $\alpha \in [0,1]$.

We also note that $\mathbf{d}^\alpha$ satisfies the linearized active equality constraints $\mathbf{c}_\mathcal{E} + A_\mathcal{E}^T \mathbf{d} = \mathbf{0}$. Thus any constraints of $QP(\mathbf{x}, \rho, J)$ for which $i \in \mathcal{E}_\circ$ are satisfied by $\mathbf{d}^\alpha$. The only remaining constraints of $QP(\mathbf{x}, \rho, J)$ come from $\mathcal{A}_\circ \backslash \mathcal{E}_\circ$. It follows from (3.10) and (3.12) that

$$c_i + \mathbf{a}_i^T \mathbf{d}^1 = c_i + \mathbf{a}_i^T \mathbf{p} + (\rho - p)\mathbf{a}_i^T \mathbf{s} \leq c_i + \mathbf{a}_i^T \mathbf{p} - (\rho - p)\varepsilon \leq 0$$

if

$$\rho \geq p + (c_i + \mathbf{a}_i^T \mathbf{p})/\varepsilon.$$

By definition of $\mathbf{p}$, the right hand side of this inequality is $O(h_{\mathcal{V}_\circ}(\mathbf{c}))$ so we can choose the constant μ in (3.7) sufficiently large so that $c_i + \mathbf{a}_i^T \mathbf{d}^1 \leq 0$, $i \in \mathcal{A}^\circ$. Thus $\mathbf{d}^1$ is feasible in $QP(\mathbf{x}, \rho, J)$ with respect to the linearized active inequality constraints, and hence to all the constraints, using results from above. Hence we have shown that there exist positive constants μ and κ such that $\mathbf{d}^1$ is feasible in $QP(\mathbf{x}, \rho, J)$ for all $\mathbf{x} \in \mathcal{N}^\circ$ and all ρ satisfying (3.7). Because we have shown above that J_+ also satisfies $\mathcal{V}_\circ \subseteq J_+ \subseteq \mathcal{V}_\circ \cup \mathcal{A}_\circ$, the same conclusions apply to $QP(\mathbf{x}, \rho, J_+)$.

Next we aim to obtain a bound on the predicted reduction Δq and hence show that (2.8), (3.8) and (3.9) hold. First we consider the line segment (3.12), and define $\phi(\alpha)$ to be the objective function of $QP(\mathbf{x}, \rho, J_+)$ evaluated at $\mathbf{d}^\alpha$. Denoting the objective function by $q(\mathbf{d}) = \frac{1}{2}\mathbf{d}^T B \mathbf{d} + \mathbf{g}^T \mathbf{d}$ where $\mathbf{g} = \sum_{i \in J_+} \mathbf{a}_i$, it follows from the chain rule that

$$\phi'(\alpha) = (\rho - p)\mathbf{s}^T \nabla q(\mathbf{d}^\alpha) = (\rho - p)\mathbf{s}^T(\mathbf{g} + B(\mathbf{p} + \alpha(\rho - p)\mathbf{s})).$$

Hence, using (3.10), bounds on B and $\mathbf{p}$, and $\rho > p$

$$\phi'(0) = (\rho - p)\mathbf{s}^T(\mathbf{g} + B\mathbf{p}) \leq (\rho - p)(\mathbf{s}^T B\mathbf{p} - \varepsilon) \leq (\rho - p)(2Mp - \varepsilon) < (\rho - p)(2M\rho - \varepsilon) \leq 0$$

if $\rho \leq \frac{1}{2}\varepsilon/M$. Now $\phi'' = (\rho - p)^2 \mathbf{s}^T B \mathbf{s} \leq 2(\rho - p)^2 M$ so

$$\phi'' + \phi'(0) \leq 2(\rho - p)^2 M + (\rho - p)(2Mp - \varepsilon) = (\rho - p)(2(\rho - p)M + 2Mp - \varepsilon) \leq 0$$

if $\rho \leq \frac{1}{2}\varepsilon/M$. In this case, applying Lemma 2, the minimum value of $\phi(\alpha)$ occurs at $\alpha = 1$ and the reduction in q satisfies $\phi(0) - \phi(1) \geq -\frac{1}{2}\phi'(0)$. After adding in a contribution for the change in q along $\mathbf{p}$, we may express

$$q(\mathbf{0}) - q(\mathbf{d}^1) \geq -\tfrac{1}{2}\phi'(0) + O(p) \geq \tfrac{1}{2}(\rho - p)(\varepsilon - \mathbf{s}^T B\mathbf{p}) + O(p) \geq \tfrac{1}{2}\rho\varepsilon + O(p).$$

Since $\mathbf{d}^1$ is feasible and $p = O(h_{\mathcal{V}_\circ}(\mathbf{c}))$, it follows that the predicted reduction (2.5) satisfies

$$\Delta q \geq \sum_{i \in J \backslash J_+} c_i + \tfrac{1}{2}\rho\varepsilon + O(h_{\mathcal{V}_\circ}(\mathbf{c})) \geq \tfrac{1}{2}\rho\varepsilon - \xi h_{\mathcal{V}_\circ}(\mathbf{c})$$

for some ξ sufficiently large and independent of ρ. Thus (3.8) is satisfied if $\rho \geq 6\xi h_{\mathcal{V}_\circ}(\mathbf{c})/\varepsilon$. This condition can be achieved by making the constant μ in (3.7) sufficiently large.

It follows from (3.5) and (3.8) that

$$\frac{\Delta f}{\Delta q} \geq 1 - \frac{\rho^2(m+1)nM}{\Delta q} \geq 1 - \frac{3\rho^2(m+1)nM}{\rho\varepsilon} = 1 - \frac{3\rho(m+1)nM}{\varepsilon}.$$

Then, if $\rho \leq (1-\sigma)\varepsilon/(3(m+1)nM)$ it follows that (2.8) holds.

Finally, we deduce from (2.8), (3.4) and (3.8) that

$$\Delta f - \gamma h_{J_\oplus}(\mathbf{c}(\mathbf{x}+\mathbf{d})) \geq \tfrac{1}{3}\sigma\rho\varepsilon - \gamma\rho^2 mnM \geq 0$$

if $\rho \leq \frac{1}{3}\sigma\varepsilon/(\gamma mnM)$. Thus we may define the constant κ in (3.7) to be the least of $\frac{1}{3}\sigma\varepsilon/(\gamma mnM)$ and the other upper bounds on ρ that are required earlier in the proof. *q.e.d.*

Now we proceed to analyse the algorithm of Figure 1. First we need a result that is similar to Lemma 6 of [3]. Here $\mathbf{x}^{(k)}$ and $B^{(k)}$ are fixed and we consider what happens to the solution of $QP(\mathbf{x}^{(k)}, \rho, J_k)$ as ρ is reduced.

Lemma 6 *Let the standard assumptions hold, then the inner iteration terminates finitely.*

Proof Consider iteration k. If $J_k = \mathcal{V}_k$ and $\mathbf{x}^{(k)}$ is a KT point of $NFP(\mathcal{V}_k)$, then $\mathbf{d} = \mathbf{0}$ is a KT point of $QP(\mathbf{x}^{(k)}, \rho, J_k)$ and the algorithm terminates. Otherwise, if the inner iteration fails to terminate, then the rule for reducing ρ ensures that $\rho \to 0$.

For the most part we may now use the proof of Lemma 5 in the case that $\mathbf{x} = \mathbf{x}^\circ = \mathbf{x}^{(k)}$ and ρ is sufficiently small. If the trust region constraint $\|\mathbf{d}\|_\infty \leq \rho$ is satisfied, we deduce as in Lemma 5 that $\mathcal{V}_\circ \subseteq J_\oplus \subseteq J_k \subseteq \mathcal{V}_k$ and $\mathcal{V}_\circ \subseteq J_+ \subseteq J_k \subseteq \mathcal{V}_k$. Since $\mathbf{x}^\circ = \mathbf{x}^{(k)}$ it follows that $\mathcal{V}_k = J_k = J_\oplus = J_+$ if the QP subproblem is to remain compatible. Thus if $\mathcal{V}_k \backslash J_k$ is not empty, then $QP(\mathbf{x}^{(k)}, \rho, J_k)$ is incompatible if ρ is sufficiently small, and the inner iteration terminates on this account. It follows from $\mathcal{V}_k = J_k$ that $h^{(k)} = 0$.

For sufficiently small ρ, it now follows from Lemma 5 and (3.8) that an f-type step is generated that satisfies the sufficient reduction condition (2.8). Also it follows from Lemma 4 that a step acceptable to the filter is generated. Thus the inner iteration terminates for sufficiently small ρ. *q.e.d.*

We are now in a position to state our main theorem. We need however to make an additional assumption. We discuss the conditions under which this assumption might hold in Section 4.

Supplementary Assumption
Under the conditions of Lemma 5, if $\mathbf{x}^{(k)}$ is in the neighbourhood $\mathcal{N}_\circ$, then the predicted reduction satisfies (3.8) for all $\rho > \kappa$.

Theorem 1 *If standard assumptions and the supplementary assumption hold, then the outcome of applying the algorithm of Figure 1 is one of the following.*

(𝔸) *A feasible point* $\mathbf{x}^{(k)}$ *of (1.1) is found.*

(𝔹) *A KT point of problem NFP*$(\mathcal{V}_k)$ *is found (*$\mathbf{d} = \mathbf{0}$ *solves* $QP(\mathbf{x}^{(k)}, \rho, J_k)$ *for some k).*

(ℂ) *There exists an accumulation point* $\mathbf{x}^\infty$ *that is feasible in (1.1).*

(𝔻) *There exists an accumulation point* $\mathbf{x}^\infty$ *that either fails to satisfy MFCQ or is a KT point of NFP*$(\mathcal{V}_\circ)$.

Proof We note in case (𝔹) that if $\mathcal{V}_k \backslash J_k$ is not empty then $\mathbf{d} = \mathbf{0}$ is not a feasible point of $QP(\mathbf{x}^{(k)}, \rho, J_k)$, which is a contradiction. Thus $\mathcal{V}_k = J_k$ and hence $\mathbf{x}^{(k)}$ is a KT point of problem $NFP(\mathcal{V}_k)$. Otherwise we need only consider the case in which none of (𝔸), (𝔹) or (ℂ) occur, and MFCQ is satisfied in case (𝔻). Because the inner loop of each iteration is finite (Lemma 6), the outer iteration sequence indexed by k is infinite. All iterates $\mathbf{x}^{(k)}$ lie in X, which is bounded, so it follows that the sequence has one or more accumulation points. Because (ℂ) does not occur, it follows that $\mathcal{V}_\infty$ is not empty.

First, we consider the case that the main sequence contains an infinite number of h-type iterations, and we consider this subsequence. For an h-type iteration, $(h^{(k)}, f^{(k)})$ is always entered into the filter at the completion of the iteration, so it follows from the corollary to Lemma 1 that $h^{(k)} \to 0$ on this subsequence, and hence $c_i^\infty = 0$ for $i \in \mathcal{J}^\perp$, where $\mathcal{J}$ is any set that occurs infinitely. It follows that $\mathcal{J}^\perp \subseteq \mathcal{V}_\infty^\perp$ and hence $\mathcal{V}_\infty \subseteq \mathcal{J}$. It also follows from (1.4) that $\mathcal{J} \subseteq \mathcal{V}_k$ and hence by continuity that

$$\mathcal{V}_\infty \subseteq \mathcal{J} \subseteq \mathcal{V}_\infty \cup \mathcal{A}_\infty. \tag{3.13}$$

It must also follow that $\tau^{(k)} \to 0$. Moreover, only h-type iterations can reset $\tau^{(k)}$, so there exists a thinner infinite subsequence on which $\tau^{(k+1)} = h^{(k)} < \tau^{(k)}$ is set. We can extract a yet thinner sequence on which the set $J_k = \mathcal{J}$ is constant. Thus we consider an accumulation point $\mathbf{x}^\infty$ and a subsequence indexed by $k \in \mathcal{S}$ of h-type iterations for which $\mathbf{x}^{(k)} \to \mathbf{x}^\infty$, $h^{(k)} \to 0$, $\tau^{(k+1)} = h^{(k)} < \tau^{(k)}$ and $J_k = \mathcal{J}$.

We now examine the proposition that $\mathbf{x}^\infty$ is not a KT point of $NFP(\mathcal{V}_\circ)$. Because MFCQ is satisfied, the vectors $\mathbf{a}_i^\infty$, $i \in \mathcal{E}_\infty$ are linearly independent, and the set defined by (3.1), (3.2) and (3.3) is not empty. For sufficiently large $k \in \mathcal{S}$ it follows that $\mathbf{x}^{(k)}$ is in the neighbourhood $\mathcal{N}^\infty$, as defined in Lemma 5. We show that this leads to a contradiction.

Lemma 5 provides conditions on ρ which ensure that $QP(\mathbf{x}^{(k)}, \rho, J_k)$ is compatible, and the step $\mathbf{d}$ obtained by solving $QP(\mathbf{x}^{(k)}, \rho, J_+)$ satisfies $\Delta f \geq \sigma \Delta q > 0$ and $f^{(k)} \geq f + \gamma h$, where f and h denote $f = f(\mathbf{x}^{(k)} + \mathbf{d})$ and $h = h(\mathbf{c}(\mathbf{x}^{(k)} + \mathbf{d}))$, respectively. This shows that the conditions for an f-type step are satisfied, and the entry (h, f) is acceptable to (not dominated by) $(h^{(k)}, f^{(k)})$. Moreover, it follows from Lemma 4 that $\mathbf{x}^{(k)} + \mathbf{d}$ is acceptable to the filter if $\rho^2 \leq \beta \tau^{(k)}/(mnM)$. Thus we deduce that if ρ satisfies

$$\mu h^{(k)} \leq \rho \leq \min \left\{ \sqrt{\frac{2\beta\tau^{(k)}}{mnM}}, \kappa \right\}, \tag{3.14}$$

then (h, f) satisfies all the conditions for an f-type iteration to occur.

Now we need to show that a value of ρ in this range will be located by the inner iteration. It follows for $k \in \mathcal{S}$ sufficiently large that $\tau^{(k)} \to 0$ and the range (3.14) becomes

$$\mu h^{(k)} \leq \rho \leq \sqrt{\frac{2\beta\tau^{(k)}}{mnM}}. \tag{3.15}$$

In the limit, because $h^{(k)} < \tau^{(k)}$, and because of the square root, the upper bound in (3.15) is more than twice the lower bound. Now consider how the inner loop of the algorithm works. Initially a value $\rho \geq \rho^\circ$ is chosen, which in the limit will be greater than the upper bound in (3.15). Then, successively halving ρ in the inner loop will eventually locate a value in the interval (3.15), or to the right of this interval. If a step is accepted for any $\rho > \kappa$ it follows by the supplementary assumption that $\Delta q \geq \frac{1}{3}\rho\varepsilon > 0$, which implies that the step is f-type. If ρ is in the interval (3.7) then we have shown that the step is f-type and is acceptable. Thus if $k \in \mathcal{S}$ is sufficiently large, an f-type iteration will result. This contradicts the fact that the subsequence is composed of h-type iterations. Thus $\mathbf{x}^\infty$ is a KT point and $\mathbb{D}$ is established in this case.

Next we consider the alternative case that the main sequence contains only a finite number of h-type iterations. Hence there exists an index K such that all iterations are f-type iterations for all $k \geq K$. It follows that $(h^{(k+1)}, f^{(k+1)})$ is always acceptable to $(h^{(k)}, f^{(k)})$, and also that $\Delta f^{(k)} \geq \sigma \Delta q^{(k)} > 0$, so that the sequence of function values $\{f^{(k)}\}$ is strictly monotonically decreasing for $k \geq K$. It therefore follows from Lemma 1 that $h^{(k)} \to 0$, and hence that any accumulation point $\mathbf{x}^\infty$ of the main sequence is a feasible point. Because $f(\mathbf{x})$ is bounded on X it also follows that $\sum_{k \geq K} \Delta f^{(k)}$ is convergent. As above, we now aim to contradict the proposition that there exists an accumulation point at which MFCQ holds that is not a KT point.

Because all iterations $k \geq K$ are f-type, no filter entries are made and so $\tau^{(k)} = \tau^{(K)}$ is constant. For sufficiently large $k \geq K$ it follows that $\mathbf{x}^{(k)}$ is in the neighbourhood $\mathcal{N}^\infty$ defined in Lemma 5. It follows as above that sufficient conditions for

accepting an f-type step are that ρ satisfies

$$\mu h^{(k)} \leq \rho \leq \min \left\{ \sqrt{\frac{2\beta\tau^{(K)}}{mnM}},\ \kappa \right\}. \tag{3.16}$$

This time the right hand side of (3.16) is a constant, $\bar{\rho}$ say ($\bar{\rho} > 0$) independent of k, whilst the left hand side converges to zero. Thus, for sufficiently large k, the upper bound must be greater than twice the lower bound. In this case, as ρ is reduced in the inner loop, either it must eventually fall within this interval or a value to the right of the interval is accepted. Hence we can guarantee that a value $\rho^{(k)} \geq \min(\frac{1}{2}\bar{\rho},\ \rho^{\circ})$ will be chosen. Using the supplementary assumption for $\rho > \kappa$, we then deduce from (2.8) and (3.8) that $\Delta f^{(k)} \geq \frac{1}{3}\sigma\varepsilon \min(\frac{1}{2}\bar{\rho},\ \rho^{\circ})$ which contradicts the fact that $\sum_{k \geq K} \Delta f^{(k)}$ is convergent. Thus $\mathbf{x}^{\infty}$ is a KT point and $\mathbb{D}$ is established in this case also. *q.e.d.*

4 Discussion

Of course, the prototype algorithm that we have outlined is flexible and may be implemented in various different ways. For example, the rule for adjusting ρ in the inner iteration could be more intricate, based partly on interpolation. The choice of an initial value of ρ for the inner iteration requires that the condition $\rho \geq \rho_{\min}$ is satisfied, but is otherwise unspecific. We envisage that in practice $\rho_{\min}$ is close to zero (say 10^{-4}) so that the effect of this restriction is negligible. Thus to a large extent these algorithms allow the more usual trust region procedure in which one may double or halve (say) the value of ρ from the previous iteration, only setting $\rho = \rho_{\min}$ if it would otherwise be less than $\rho_{\min}$. We have successfully used such ideas in our codes.

An important issue is that of how to specify the matrices $B^{(k)}$ that are used to define the QP subproblems. There is the issue of whether to use exact Hessians and what multiplier estimates to use. If quasi-Newton updates to the Hessian are used then there is the question of which update formula to use. Also it can be important to pay attention to the asymptotic behaviour of the algorithm, to ensure that the second order convergence property of the SQP iteration is not compromised. It is not yet clear for NLP filter algorithms how best to do this, even though second order convergence is almost invariably observed for regular problems. Also there is the question of whether it is advantageous to use SOC steps. All these issues must be addressed when the algorithms that we suggest here are implemented.

There is one idea that we do feel is useful, in regard to that step in the algorithm in which $QP(\mathbf{x}^{(k)}, \rho, J_k)$ is inconsistent and some constraints in $\mathcal{V}_k \backslash J_k$ are relaxed. One possible way to proceed is to relax all constraints in $\mathcal{V}_k \backslash J_k$, but this can be disadvantageous because it creates a QP subproblem with a large null space, that can be expensive to solve. An effective way of choosing J_{k+1} that usually relaxes only a few constraints is to make use of a Phase 1 LP solver that is based on minimizing the

existing sum of infeasibilities, until either a feasible point is obtained or no further improvement is possible. Only those constraints that are infeasible in the solution given by the Phase 1 solver are relaxed. Some care has to be taken that linear constraints do not go infeasible in the Phase 1 solver. Also it is necessary to ensure that the resulting (h, f) pair is acceptable to the filter. If not, further constraints are relaxed, one at a time, starting with the most infeasible, until an acceptable (h, f) pair is located.

In regard to the calculation of a set J_+ such that $\mathbf{d}$ and J_+ conform, one idea has already been described in Section 2. Another way to proceed is the following. Initially, in $QP(\mathbf{x}^{(k)}, \rho, J_k)$, there are no constraints on $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d}$ for $i \in J_k$. First we introduce the reverse inequality $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} \geq 0$ into the QP subproblem for all such constraints, and re-solve the new QP subproblem. If all the inequalities $c_i^{(k)} + \mathbf{a}_i^{(k)T}\mathbf{d} \geq 0$ $i \in J_k$ are inactive then we have a conforming solution. Otherwise we relax any reverse constraints that have become active and repeat the QP solution. This process is repeated until no reverse inequalities are active, in which case a conforming solution has been obtained. A feature of this method, if global solutions of the QP are calculated, is that the value of Δq increases monotonically as ρ is increased. The opposite is observed in practice for the technique described in Section 2. An advantage of the technique of Section 2 is that the feasibility of $QP(\mathbf{x}^{(k)}, \rho, J_k)$ is immediately determined on the first step of the process.

Another way to compute a set J_+ such that $\mathbf{d}$ and J_+ conform is to solve the subproblem

$$\begin{array}{ll} \underset{\mathbf{d}\in\mathbb{R}^n}{\text{minimize}} & \frac{1}{2}\mathbf{d}^T B^{(k)}\mathbf{d} + \sum_{i\in J_k}(c_i + \mathbf{a}_i^T\mathbf{d})^+ \\ \text{subject to} & \|\mathbf{d}\|_\infty \leq \rho \end{array} \tag{4.1}$$

for each value of ρ on the inner iteration. This has the advantage that if global solutions of the QP are calculated then an optimum choice of Δq is made, and the supplementary assumption can be dispensed with. However, this subproblem is not a standard QP problem, so implementation is much less convenient. Also, on many iterations the set J_+ that would be obtained is no different from that obtained by the method of Section 2, for example when $J_+ = J_k$ as often happens. Thus it is likely that there is little to be gained by using this technique.

We now present some observations on whether the assumptions under which Theorem 1 holds are likely to be valid. For the Standard Assumptions, the only case which is likely to cause difficulties is if an a-priori bound on $\|B^{(k)}\|$ is not available. This is likely to be correlated with the issue of whether or not any multiplier estimates are bounded, since these are used directly if exact Hessians are calculated, or indirectly in the updating formula if quasi-Newton updates are used. But we have already observed that if MFCQ holds, then multiplier estimates are bounded. Thus it is likely in practice that $\|B^{(k)}\|$ will be bounded in this case.

The validity of the Supplementary Assumption is less easy to guarantee. An obvious difficulty arises when the matrix $B^{(k)}$ is indefinite, and local but not global solutions are calculated by the QP solver. However, even if global solutions are calculated, for example when the matrices $B^{(k)}$ are positive semi-definite, then all is not straightforward. This is because the technique for finding a solution in which $\mathbf{d}$ and J_+ conform, does not obviously (to us) lead to a proof that Δq is a monotonically increasing function of ρ, even if global solutions of the QP are assumed. (Such an argument has been used in other filter-type proofs). On the other hand, as mentioned above, the set J_+ will often be the same as that calculated by (4.1), in which case (3.8) can be deduced from an assumption that global solutions of the QP are satisfied. Moreover, difficulties do not seem to arise in practice with trust region methods for other types of problem due to the calculation of local but non-global solutions of the QP subproblems. Thus it seems very unlikely that condition (3.8) will fail to hold in practice, near a non-KT point.

We have implemented a feasibility restoration technique that is based on the algorithm given in this paper, and is consistent with the convergence theory. Numerical results are comparable to those described in [2]. A quasi-Newton version of the technique is also under development. We hope to report on the performance of these codes in a future paper.

5 References

[1] Chin C.M. and Fletcher R. (2001), On the Global Convergence of an SLP-filter algorithm that takes EQP steps, Dundee University, Dept. of Mathematics, Report NA/199.

[2] Fletcher R. and Leyffer S. (1997), Nonlinear Programming Without a Penalty Function, Dundee University, Dept. of Mathematics, Report NA/171, to appear in Mathematical Programming.

[3] Fletcher R., Leyffer S. and Toint Ph.L. (2000), On the Global Convergence of a Filter–SQP Algorithm, Dundee University, Dept. of Mathematics, Report NA/197, to appear in SIAM Journal of Optimization.

[4] Gauvin J. (1977), A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming, *Mathematical Programming*, **12**, pp. 136–138.

[5] Powell M.J.D. (1970), A hybrid method for nonlinear equations, in *Numerical Methods for Nonlinear Algebraic Equations*, (Ed. P. Rabinowitz), Gordon and Breach, London.

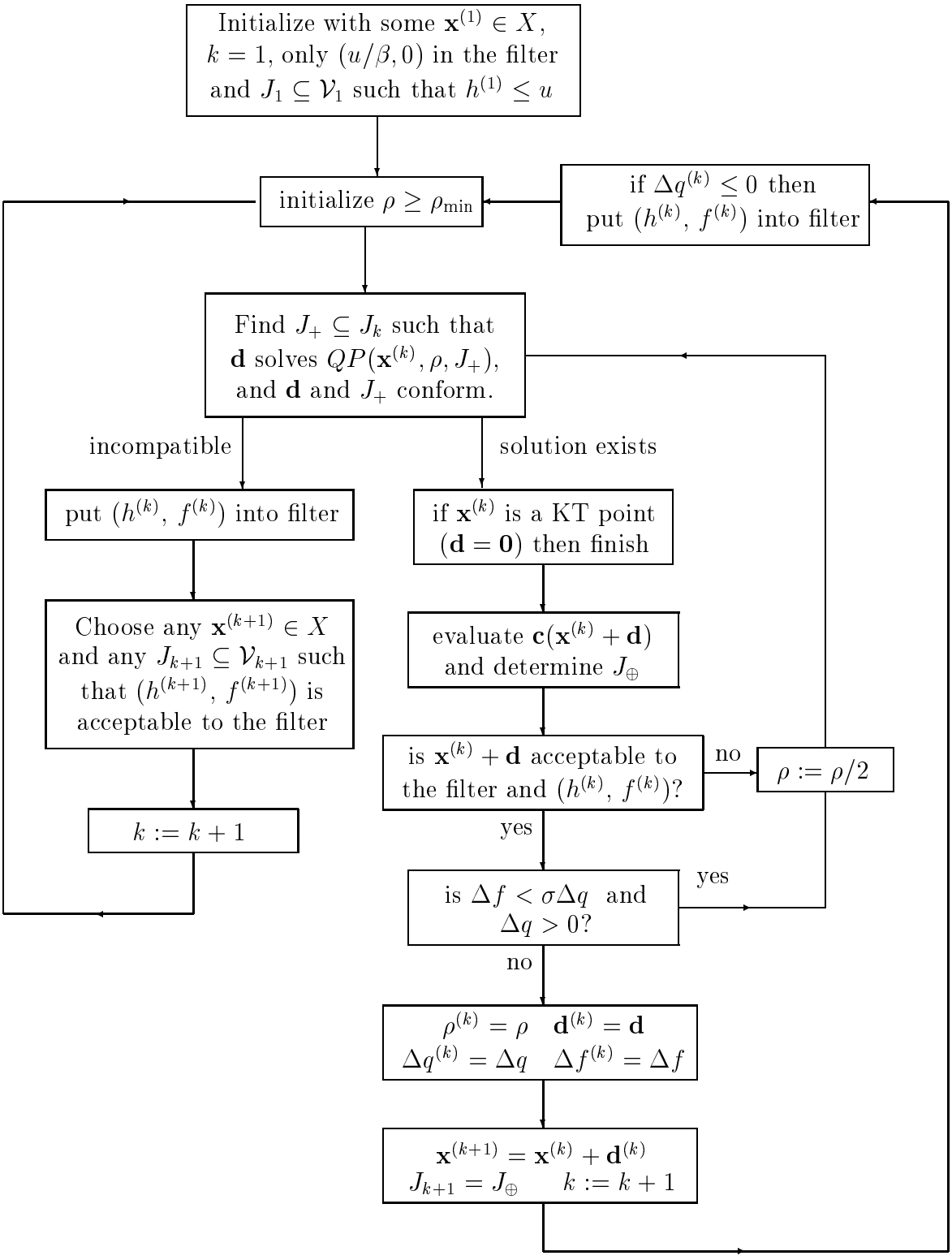

Figure 1: An SQP Filter Algorithm